

第六项议程

2018–19 年度评估报告

文件的目 的

本文件向理事会提交有关国际劳工组织评估办公室 2018-19 年度的工作进度报告，这项进度对照其基于结果的《2018-21 年评估战略》中的指标和目标加以衡量。本文件还采用一整套绩效指标，根据本文件所涉期间对多项目评估的多元研究，研讨国际劳工组织整体有效性。现请理事会核可本报告所载建议，这些建议将列入国际劳工组织建议落实工作的持续计划，还请理事会确认 2019-21 年评估工作计划的优先事项(见决定草案：第 74 段)。

相关的战略目标：全部。

主要相关成果/跨领域政策驱动因素：扶持性成果 B：本组织的有效和高效治理。

政策影响：本报告建议将具有政策影响。

法律影响：无。

财政影响：现有工作人员的时间。

需采取的后续行动：经批准的建议将被插入用以跟踪后续行动的持续行动计划。

作者单位：评估办公室(EVAL)。

相关文件：理事会文件 GB.334/PFA/6；GB.332/PFA/8；GB.331/PFA/8。

目 录

	页 次
导 言	1
第一部分. 实施国际劳工组织评估战略	2
实现关键里程碑方面取得的进展.....	2
成果 1: 增强评估能力和体系, 以利实践和使用	2
子成果 1.1 – 依据《评估政策》要求及时开展的评估活动	2
子成果 1.2 – 加强各区域与各部门工作人员的评估能力	4
子成果 1.3 – 三方成员以回应可持续发展目标的方式参与 监测和评估《体面劳动国别计划》以及发展合作活动	6
子成果 1.4 – 评估成为《体面劳动国别计划》和发展合作 活动的一部分, 侧重可持续发展目标	7
子成果 1.5 – 加强各区域及部门的能力, 纳入并使用评估	8
成果 2: 通过使用可信度更高且更优质的评估(独立性、可信度、 有用性)提高评估价值.....	9
子成果 2.1 – 使用战略集群评估, 更有效地收集评估信息	9
子成果 2.2 – 提高内部评估、分散评估及集中评估的质量	10
子成果 2.3 – 为有效政策干预实行知识建设开展可信度高 的影响评估	14
子成果 2.4 – 更符合国际劳工组织权责和包括可持续发展 目标在内的具体情况的评估框架	15
成果 3: 加强评估结论和建议知识库	15
子成果 3.1 – 借助 i-eval Discovery 应用程序增强评估信息 的可获取性及知名度	15
子成果 3.2 – 加强对评估结论的定向交流	16
子成果 3.3 – 三方成员及管理层更好地将评估结论及建议 用于管理及决策过程	17
子成果 3.4 – 通过进一步分析评估结论及结果, 利用评估 满足战略知识的要求	23

	页 次
第二部分. 国际劳工组织有效性及结果评估.....	25
1. 制定国际劳工组织绩效在国家及全球两级更全面的验证程序： 进展受限的内外部因素	25
2. 国际劳工组织工作的整体有效性.....	25
3. 促进组织学习：有利于成功项目效力的主要驱动因素	29
决定草案	30
附 录. 执行《2017-18 年度评估报告》中所载已批准建议的行动计划.....	31

导 言

1. 2019 年是国际劳工组织百年纪念之年，各种高级别论坛和声明都重复提及基于结果管理的重要性。《国际劳工组织关于劳动世界的未来百年宣言》强调知识管理对进一步提高循证政策咨询质量的重要意义。知识管理和评估密不可分，都有助于协作学习。国际劳工组织评估办公室(EVAL)在一百周年之际编写的一份案例研究记录了国际劳工组织六十年来评估的历史演变。该研究将过去十年描述为一个变革时期，其特点是评估职能独立和现代化，在此期间“进行了重大的组织变革，以确保其结构独立性……采取大胆的措施协调程序、提高合规性、质量和使用情况，同时，通过基于网络的指导和工具使操作现代化”。¹ 这项声明全面概括国际劳工组织目前的评估状况，本年度报告将对此进行进一步探讨和说明。
2. 本报告的第一部分详细介绍评估职能在实施评估战略方面取得的进展。国际劳工组织内部，对监测、评估和衡量有效性的兴趣日益增加(例如，基于结果管理工作队)²，这为有效执行评估战略创造了有利环境。重视优质设计(包括变革理论)、监测和报告以及制定新一代《体面劳动国别计划》(DWCPs)已取得巨大的发展势头。然而，这些变化产生影响尚需假以时日，但终将影响评估战略实施的节奏。
3. 现在正在考虑根据地理位置和专题对项目 and 计划活动进行战略集群评估，但这些评估仍有待在国际劳工组织内部扎根。已经制定并在逐步采用确保更符合国际劳工组织任务和具体情况(例如，三方性、社会对话和标准)的新的准则，及其方针、方法和框架。多年来推进并致力于能力建设、指导和知识体系，现已开始出现成果，具体体现在评估报告质量得以提高。2011 年开始的外部团队对评估报告的事后质量审查最终表明报告质量有所提高，并更好地涵盖了性别问题。评估办公室历时多年开发的一套知识管理工具在内部和外部都被视为鼓励采用评价结果的创新手段。为国际劳工组织工作人员、三方成员和其他利益攸关方提供的与评估相关的能力建设培训计划也继续蓬勃发展，从而促进整个组织评估网络和评估文化的发展。
4. 本报告的第二部分评估国际劳工组织整体有效性。基于对采用一整套绩效指标评估发展合作项目所做的多元研究，评估办公室继续借鉴以往的经验衡量体面劳动结果(有效性评估)。对国际劳工组织整体有效性的最新评估显示，项目设计质量、监测

¹ 国际劳工组织： [国际劳工组织评估简史](#)，评估办公室(日内瓦，2019)。

² 根据其职权范围，基于结果管理工作队的任务是：(i) 审查国际劳工组织结果框架和国际劳工组织战略预算程序以及相关报告程序；(ii) 在全劳工局(包括在发展合作项目中)修订和加强国际劳工组织对“变革理论”方法的使用，以进行各个级别的成果测量、因果关系分析和对影响和可持续性的评价，从而改进绩效报告和问责制；(iii) 加强各级工作计划和监测以及框架及系统评估，以确保结果信息流、汇总及可评估性，以及(iv) 加强劳工局支持成果管理现有体系的能力(综合资源信息系统，信息汇总等)。

及报告制度方面仍存在不足，并且项目过于雄心勃勃，体现在实现承诺成果的资源不足。为了促进组织学习，该研究确定促使项目取得成功的有利因素，以供系统地采用，并仔细审查由经常预算补充账户(RBSA)供资的项目样本的绩效。

5. 由于该有效性评估为追溯性评估，因此近年来采用的编程实践所产生的变化尚未显示出预期绩效成果。另一局限性在于，评估方法仅通过一个代理指标(项目评估)做到接近总体效力，因此不能真正全面洞察组织有效性。正如去年的年度评估报告所解释的那样，最佳局面是评估体系可以全面涵盖并验证所有《体面劳动国别计划》与计划和预算成果的结果，一个计划周期为四年。然而，这将要求国际劳工组织在进行项目评估的方式上做出重大改变。这一体系的组成部分正在逐步到位，但速度取决于扶持性组织环境可能并将发展的速度。

第一部分. 实施国际劳工组织评估战略

实现关键里程碑方面取得的进展

6. 本年度评估报告的第一部分对照国际劳工组织基于结果的《2018-21 年评估战略》，对国际劳工组织评估职能的绩效进行评价。
7. 2017-18 年度评估报告是 2018 年中期报告，今年的报告则反映 2018-19 两年期取得的进展，因此可以对所取得的成就进行明确评价。总体而言，两年期的进展有令人满意之处：总共 19 个两年期里程碑中，9 个达到或超过了目标，9 个部分达到了目标，1 个没有取得任何进展。本报告中关于评估战略的报告符合并补充了国际劳工组织关于扶持性成果 B(本组织的有效和高效管理)的计划执行报告。
8. 稳固的有利成果环境、评估文化以及评估职能的独立性至关重要，评估职能只有在这种条件下才可以不受任何因素的影响，才可以对国际劳工组织的计划有效性作出可信的评价。多年以来，评估办公室成功建立了必要的结构和程序，以避免评估职能的可信性受到削弱。

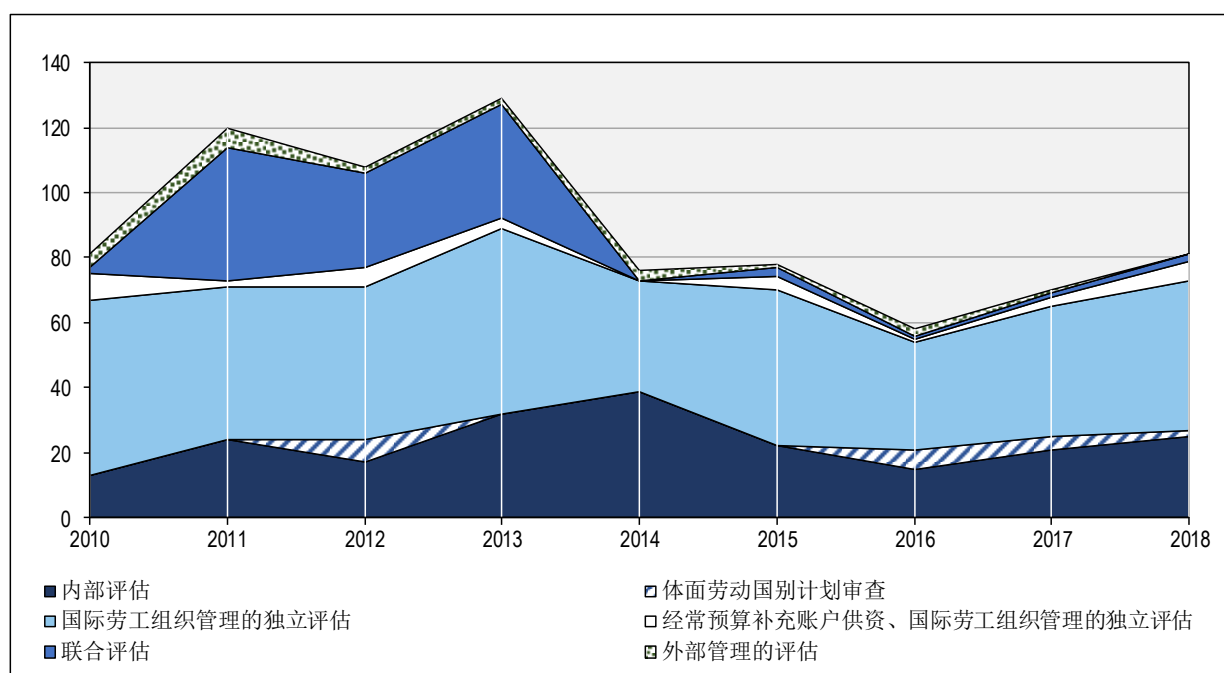
成果 1： 增强评估能力和体系，以利实践和使用

子成果 1.1 – 依据《评估政策》要求及时开展的评估活动

指标 1.1: 及时完成所有强制性评估，供管理部门、三方成员和捐助者使用。	基线: 独立评估覆盖率达到 90%，内部评估覆盖率达到 33%。	状态 部分达成。
两年期里程碑 (2018–19): 两年期末及时完成 95%独立评估及 50%内部评估(目标)。		
最终目标 (2020–21): 2021 年底前，及时完成 95%独立评估和 75%内部评估，以影响决策。		

9. 最佳成果管理和知识管理取决于可靠、可信的评估结果。评估办公室向公众开放的 *i-eval Discovery* 网络应用程序说明了所有需要评估的项目(见子成果 3.1 下的两年期里程碑)。
10. 2018 年, 总共 95 项评价中(见图 1)– 共完成了 54 项独立评估,³ 比上一年增加了 13 项。2017-2018 报告年度除一项独立评估外, 其他所有评估均按要求进行。⁴ 因此, 独立评估实现了 2018-2019 两年期里程碑, 完成率为 99%。
11. 然而, 内部评估的数字不那么乐观, 提交率低于预期。2017 年, 评估办公室启动了内部评估认证计划(IECP), 以帮助国际劳工组织工作人员进行能力建设, 从而进行内部评估并提高完成率(见指标 1.2.1)。《2017-2018 年度评估报告》中内部评价的完成率为 36%。加上 2018-2019 年 46% 的报告数字, 两年期的完成率为 41%。因此, 劳工局没有达到 50% 的目标, 但是离目标接近。作为《2017-2018 年度评估报告》中提出并经理事会认可的建议后续行动, 自我评价部分现已纳入进度报告定稿中,⁵ 以努力提高报告效率。

图 1. 2010-18 年按类型列出的评估数目



³ 这些项目包括: 一项评估性评价, 六项经常预算补充账户评价(或含有经常预算补充账户成分的项目), 三项集群评估和两项联合评估。无外部评估。

⁴ 如《2017-18 年度评估报告》所述。

⁵ 价值小于 500,000 美元的项目需要进行自我评估, 并应存储在捐助方报告平台中。

选择高级评估专题用于战略用途

12. 每年都制定用于战略用途的高级评估专题滚动工作计划。制定过程涉及与三方代表协商、与评估咨询委员会讨论，还要对国际劳工组织的管理文件进行彻底的审查，以明确间接或者直接提到的评估需要的决定。评估办公室权衡所收到的有评估需要的信息，以确保长期未被评估的专题得到适当关注(表 1)。

表 1. 2019-23 年高级评估专题摘要

年	机构或成果级别	成果级别	体面劳动 国别计划 ¹	三方成员意见
2023	农村经济 ²	不可接受的劳动形式 ³	阿拉伯国家	普遍支持，但对新的计划和预算结构对拟议专题的影响表示关切。还有人选定 2023 年的议题为时过早表示关切。
2022	劳工统计开发与利用 ⁴	国际劳工标准实施 ⁵	欧洲	普遍支持，但对评估国际劳工标准的适当性表示关切。还有人选定 2022 年的专题为时过早表示关切。
2021	性别平等行动计划 ⁶	推行公平有效的劳务移民政策 ⁷	亚洲	再次确认，但有人对既注重移徙又注重公平征聘表示关切，因为大会决议只要求评估促进公平征聘的工作。
2020	研究与知识管理 ⁸	促进可持续发展企业	美洲	再次确认，但有人对在同一次理事会会议上对评估与国际劳工组织《研究议程》的讨论之间的协调表示关切。

¹《体面劳动国别计划》评估在地区间轮换进行。²以前未评估过。³至少在两个两年期内没有作为成果目标进行评估。⁴机构性：以前未评估过；根据以前的协商提供的信息选定。⁵成果目标：至少在两个两年期内没有作为成果目标进行评估。根据以前的协商提供的信息选定。⁶机构性：理事会已批准的行动计划部分到期。⁷成果：2017 年 6 月通过的国际劳工大会关于公平和有效劳务移民管理的决议要求对促进公平招聘的工作进行高级别评估。该评估最初安排在 2019 年进行，后推迟至 2021 年。⁸机构性：以前未评估过；根据以前的协商提供的信息选定。

子成果 1.2 – 加强各区域与各部门工作人员的评估能力

指标 1.2.1: 提高了国际劳工组织工作人员的评估能力。	基线: 2017 年底，77 名工作人员被认证为考核管理员，两名被认证为内部评估认证计划的一部分。2016-17 年度颁发了八份评估管理员认证计划证书。	状态 超额完成。
两年期里程碑 (2018–19): 至少还要有 30 名国际劳工组织的工作人员被认证为评估管理员和内部评估人员。		
最终目标 (2020–21): 2021 年底前，至少还有 120 名国际劳工组织的工作人员要被认证为评估管理员和内部评估人员。		

13. 评估办公室主要通过评估管理员认证计划(EMCP)和内部评估认证计划(IECP)实现这一里程碑。除了这些计划以外，各地区偶尔为其工作人员举办针对监测和评估的一般性培训。

14. 2018 年, 7 名评估管理员认证计划参训人员和 11 名内部评估认证计划参训人员完成了所有严格的认证要求。2019 年, 又有 20 名工作人员被认证为评估管理员, 另有 10 名工作人员被认证为内部评估员。2018-2019 年共有 48 名参训人员获得认证, 超过了两年里程碑 30 名的目标(表 2)。评估管理员认证计划课程于 2013 年启动, 对于一些证书持有人来说, 他们接受培训已经六年。为了帮助获得认证的工作人员更新他们的能力并获得更高水平的认证, 评估办公室正在开发一门评估管理员认证计划高级课程, 将于 2019 年底推出。

表 2. 2019 年已接受评估培训的国际劳工组织官员¹

培训类型	非洲	美洲	阿拉伯国家	亚太地区	欧洲	总部	总和
普通检测及评估	1	4	23	8	25	0	61
评估管理技能	20 ²	0	2	16	0	0	38
内部评估技能	4	0	0	2	4	3	13
总和	25	4	25	26	29	3	112

¹ 培训只是认证要求之一; 还必须成功完成实习。

² 预计数字, 因为 2019 年第四季度还有培训。只有至少超过一天的培训计划才包括普通监测和评估培训。

指标 1.2.2: 国际劳工组织评估网络的运作基于明确的任务和岗位描述。	基线: 目前, 评估网络职能(部门级和评估管理员级)在自愿的基础上履行。	状态 部分达成。
两年期里程碑 (2018–19): 评估职责包括在部门评估协调员的岗位描述中, 被认证的评估管理员在其绩效考核中接受标准化评估。		
最终目标 (2020–21): 2021年底前, 功能齐全的评估网络牢固地嵌入相关的区域和部门职能, 并配有适当的资源和激励措施。		

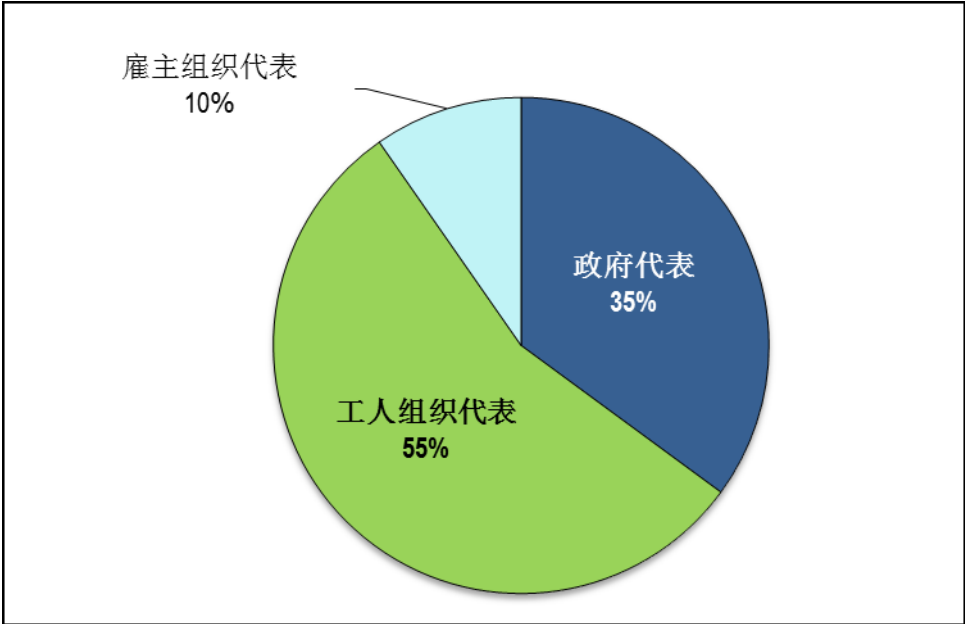
15. 有必要加强针对主动参与评估网络(即部门评估联络点, 评估管理员和内部评估员)的国际劳工组织同事的激励结构。评估办公室开发了具体的、可测量的、可实现的, 现实的和及时的(SMART)输出和绩效衡量标准, 以便可以纳入国际劳工组织的绩效评估体系。该系统包括一个附带指标和目标的模型输出库, 可供部门评估联络点、评估管理员和内部评估员进行周期讨论使用。此外, 还提出了建议, 要求开发可纳入岗位描述的标准评估任务。一旦采用, 不仅国际劳工组织同事主动参与评估网络的工作可以得到恰如其分地认可, 他们还可以在其绩效评估中接受标准化评估。该材料已与人力资源开发部共享, 正等待实施。

子成果 1.3 – 三方成员以回应可持续发展目标的方式参与监测和评估《体面劳动国别计划》以及发展合作活动

指标 1.3： 相关监测和评估培训被纳入为三方成员举办的培训和能力建设计划，以提高他们的评估参与率。	基线： 2010-17 年期间，1,052 名三方成员接受了培训，其中 124 名于 2016 年接受了培训。	状态 已达成。
两年期里程碑 (2018–19)： 开发对可持续发展目标予以回应的评估培训和能力建设模块，并纳入计划，涵盖所有三方成员。		
最终目标 (2020–21)： 2021年底前，作为扩大的评估办公室和国际劳工组织范围培训计划的一部分，至少有 150 名三方成员(三个组比例相等)接受过量身定制的评估培训。		

16. 2018 年 11 月，评估办公室为国际劳工组织三方成员启动了一个关于评估可持续发展目标时代下的《体面劳动议程》培训计划。有些培训计划由各部门和区域/国家办事处合作开展(图 2)。⁶ 2018-19 两年期内，共有 134 名来自政府、雇主和工人组织的代表接受了培训。关于如何更积极地参与国家和地方各级《体面劳动议程》和可持续发展目标的规划和评估的三方讨论，为自愿国别审查提供了信息。

图 2. 已接受评估可持续发展目标时代下的《体面劳动议程》培训的国际劳工组织三方代表



⁶ 阿根廷(布宜诺斯艾利斯，2018 年 11 月)、黎巴嫩(贝鲁特，2019 年 4 月)、国际培训中心-国际劳工组织(都灵，2019 年 4 月)、阿塞拜疆(巴库，2019 年 5 月)和印度尼西亚(雅加达，2019 年 5 月)。

17. 评估办公室继续与其他部门、区域及国际劳工组织国际培训中心(ITC-ILO)合作, 将评估培训模块纳入能力建设计划。评估办公室还在探讨组织区域研讨会的可能性, 以促进拉丁美洲和加勒比区域三方成员之间就这一专题交流良好做法和学习。需要更加重视专门针对雇主的培训, 以实现接受培训的三方成员比例相等的目标。

**子成果 1.4 – 评估成为《体面劳动国别计划》
和发展合作活动的一部分, 侧重
可持续发展目标**

指标 1.4: 具有完善的评估程序和机制、定期与三方成员接触以满足监测和评估要求的《体面劳动国别计划》和发展合作项目的数量。	基线: 尚未确定基线。	状态
两年期里程碑 (2018–19): 开发并试运行了《体面劳动国别计划》的样本, 以确保《体面劳动国别计划》和项目具有机制(诊断工具), 从而对其可评估性、对可持续发展目标的回应性以及三方成员参与监测和评估的程度进行评估。		部分达成。
最终目标 (2020–21): 到 2021 年底, 75%的《体面劳动国别计划》和发展合作项目都建立完善的机制, 可以评价其可评估性、可持续发展目标回应性以及三方成员参与监测和评估的水平。		

18. 2018 年 2 月, 评估办公室开发了可评价性诊断工具,⁷ 以支持对可持续发展目标敏感的监测以及《体面劳动国别计划》评估的强化规划。2018 年, 对两次《体面劳动国别计划》设计演习进行了评价性诊断。2019 年, 又在另外三个国家进行了测试。⁸ 正在与战略规划和管理部合作汇编经验教训, 以确保回应可持续发展目标的《体面劳动国别计划》设计, 确保更多三方成员参与, 包括随后的监测和评估。
- 《联合国可持续发展合作框架》背景下的最新进展将进行一些调整。评估办公室是联合国评价小组评价和可持续发展目标工作组的共同协调员, 在该工作组中, 评估办公室和他们分享经验, 并随时了解最新发展状况。

⁷评估工具从以下三个维度提供了国家一级的分析: (i)体面劳动国别计划的设计、实施及报告是否足以说明国际劳工组织对可持续发展目标在国家一级的贡献; (ii) 国家一级监测及评估是否具备实施、监测及评估可持续发展目标的条件(侧重体面劳动); (iii) 三方成员能力需要以及有效宣传及参与可持续发展目标/体面劳动监测与评估进程所面临的挑战。

⁸ 布隆迪、伊拉克和苏里南的体面劳动国别计划。

子成果 1.5 – 加强各区域及部门的能力，纳入并使用评估

指标 1.5: 除体系强制要求以外，各区域及部门采取与评估有关的举措。	基线: 2015 年度评估报告以来尚未系统记录过此类举措实例及其使用。	状态
两年期里程碑 (2018–19): 系统记录此类举措，形成基于大型计划或旗舰计划经验的良好做法。		部分达成。
最终目标 (2020–21): 2021 年底前，将建立一个按部门和区域分列的、对倡议进行定量和定性记录的系统程序，以显示逐步增长和附加值。		

19. 各地区和部门越来越积极地采取超出强制性要求的与评估相关的举措，并已建立这些举措的系统文件。表 3 是评估办公室所发起、激励或支持的举措的非详尽概览。例如，就业促进和平与复原力小组编写了一本综合手册《如何设计、监测和评估就业促进和平与复原力计划中的建设和平成果》，其中纳入了评估办公室提供的信息，也纳入了针对国际劳工组织在脆弱国家工作的专题评价中的信息。⁹ 评估办公室还支持为一个进行中的、由荷兰出资的多伙伴合作项目开发方案监测和评估框架，该项目涉及阿拉伯国家和北非的难民和收容社区的就业、社会保障和教育问题。此外，零伤亡愿景基金及其合作伙伴德国国际合作机构(GIZ)与评估办公室合作开发了变革理论 — 全面监控计划和综合评估矩阵，以支持全球计划。最后，支持项目和计划的监测、报告和评价的正式员工人数增加到近 30 人，由此可以看出人们日益认识到将评价纳入主流的重要性。最后，支持项目和方案的监测、报告和评价的专职工作人员人数有所增加(至将近 30 人)，这也体现出人们越来越认识到主流化评估的重要性。

表 3. 2019 年已报告的非强制评估举措之非详尽概览

评估举措种类	各部门和地区评估举措的数目
多元研究/综合评论	全球 1，非洲 1，阿拉伯国家 1，亚洲和太平洋地区 2，拉丁美洲 1
专题评估	阿拉伯国家 1，亚洲和太平洋地区 1，拉丁美洲 1
影响评估	全球 1，非洲 1，阿拉伯国家 1
综合监测和评估方针	全球 2，阿拉伯国家 1
知识管理和评估交流	非洲 1，拉丁美洲 1
监测和评估能力建设	全球 2，阿拉伯国家 1，亚洲和太平洋地区 2，拉丁美洲 2
国家政策和计划评估	全球 2，非洲 1

⁹ 国际劳工组织：对国际劳工组织在冲突后、脆弱和受灾难影响的国家的工作进行独立专题评估：过去、现在和将来 (日内瓦，2015 年)。

成果 2： 通过使用可信度更高且更优质的评估(独立性、可信度、有用性)提高评估价值

子成果 2.1 – 使用战略集群评估，更有效地收集评估信息

指标 2.1： 战略集群评估成为大部分计划与项目的模式。	基线： 目前尚无现成的、记录详尽的过程或程序可用于进行发展合作项目的战略集群评估。	状态
两年期里程碑 (2018–19)： 为战略集群评估开发方法和程序，包括开发募集评估资金的模式，并至少在五个项目中试运行。		符合进度。
最终目标 (2020–21)： 2021 年底前，将有一个由相当人数的捐助者(25%)批准的战略集群评估程序。		

20. 除了潜在增效之外，将多个评估合并为一个评估的显著优势是可以分析在多个国家实施的类似项目(例如，在相似专题领域)的共性和差异。这有利于确定关键的成功因素和潜在风险，从而为将来和正在执行的干预措施提供有价值的信息。
21. 评估办公室制定了两阶段方针以支持战略集群评估的发展。第一阶段包括评述集群的经验和可能出现的情况，第二阶段旨在制定程序指南以在实践中推广。评述阶段进行的研究表明，集群有很好的潜力从战略上把项目、组织成就和影响联系起来。在没有充分认识到效益的情况下，或者由于现行政策的原因，许多捐助者仍然赞成按项目进行区分评估的方法。该研究强调，集群评估的发展和推广将需要加强评估技术、方法和程序，包括改变劳工局向捐助方和合作伙伴介绍该备选方案的做法。
22. 研究表明，与捐助方和合作伙伴的讨论目前尚未系统地传达《2017 年评估政策》的含义，其中包括为产生更大的战略影响，审查集群评估选项和募集评估资金的要求。最后，该研究警告，成本效益可能是采取集群评估的依据，但并非唯一的依据。集群评估可能不会即刻降低评估费用，因为需要正确认识到缺少评价所需募集资金的复杂性以及额外交易成本。

建议 1： 在讨论协议时，将评估战略分组作为所有捐助方的优先选择，并利用评估职能的专家意见解释符合评估政策的益处和各种选择退出的情况。

23. 作为第一阶段的一部分，已确定了五个试点项目，测试可能进行的集群评估的分类(表 4)。第二阶段正在进行，关于如何规划集群评估的指导说明预计于 2019 年底出台。

表 4. 可能进行的已确定归类的集群评估试点

	集群类型	可能进行的试点	范围和重点
地 域 (国家/区域)	《体面劳动国别计划》 或类似国家框架(国家计 划评价)	危地马拉作为捐助者成 果供资框架的一部分	有助于实现《体面劳动国别计划》 /国家计划成果的所有项目；对 “联合国可持续发展合作框架”的 投入/连接
	更广泛/跨领域国际劳工 组织干预	“国际劳工组织一体化” 埃塞俄比亚(“国际劳工 组织一体化”协调工作 评估)	处于各个阶段的多个项目
专 题	部门(例如服装部门，农 业部门)	孟加拉国成衣计划	一个国家某个部门的所有项目
	基于计划和预算成果	捐助方—国际劳工组织 伙伴计划	有助于实现类似目标/成果的项目
	全球计划	零伤亡愿景基金	不同国家使用的相似模型

24. 促进对评价进行集群以获得更大战略影响的方法还契合联合国正在进行的改革讨论，以及加强独立的全系统联合评价的努力。评估办公室积极参与联合国评估小组的工作，为全球范围内全系统独立评估制定路线图，并在选定的国家试行新一代的“联合国可持续发展合作框架”评估。

子成果 2.2 – 提高内部评估、分散评估及集中评估的质量

指标 2.2.1: 所有发展合作项目的评估都要遵守经合组织及联合国评价小组的规范和标准，并满足国际劳工组织特定任务及学习需要。	基线: 2015-17 年外部质量评估显示 90% 的发展合作项目评估符合质量标准要求。	状态 已达成。
	两年期里程碑 (2018–19): 更新指导方针，在保证质量的同时纳入新的反映国际劳工组织特定任务的评估模式。	
	最终目标 (2020–21): 2021 年底前，外部质量评价确认 95% 的发展合作项目评估要符合经合组织及联合国评价小组的标准。	

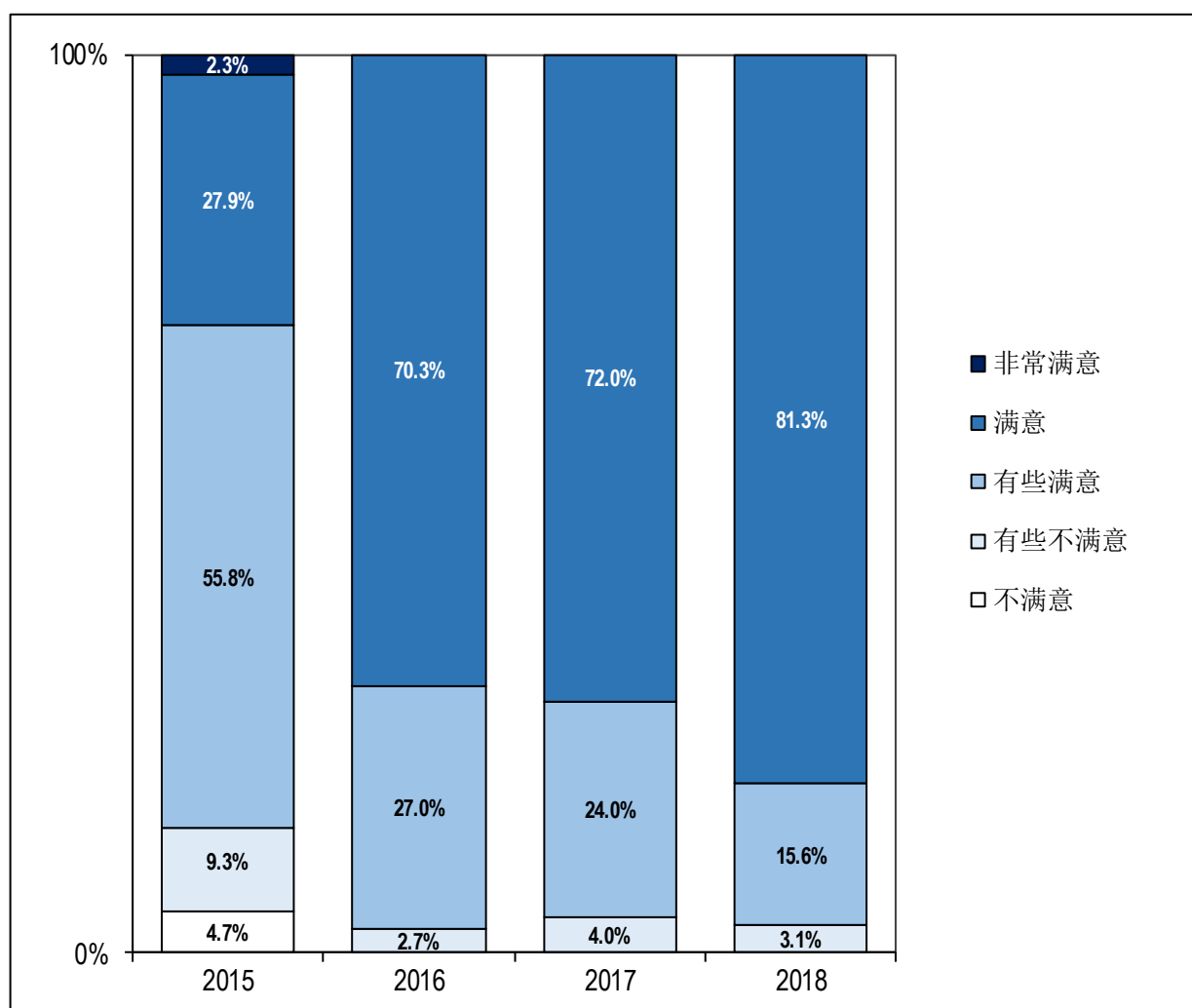
25. 评估办公室使用分层审批流程，其中涉及评估管理员，部门评估联络人，区域评估人员和高级评估人员，这些人员为项目评估提供实时质量保证。该过程由外部评估员进行的事后质量评估加以补充。¹⁰ 最新汇总结果显示，总体而言，2017-18 年度通

¹⁰ 2008 年以来，共进行了八轮独立评估质量评价。

过评价的评估报告中有 98% 获得的评级等于或高于“令人满意”，超过了达到 90% 的目标。

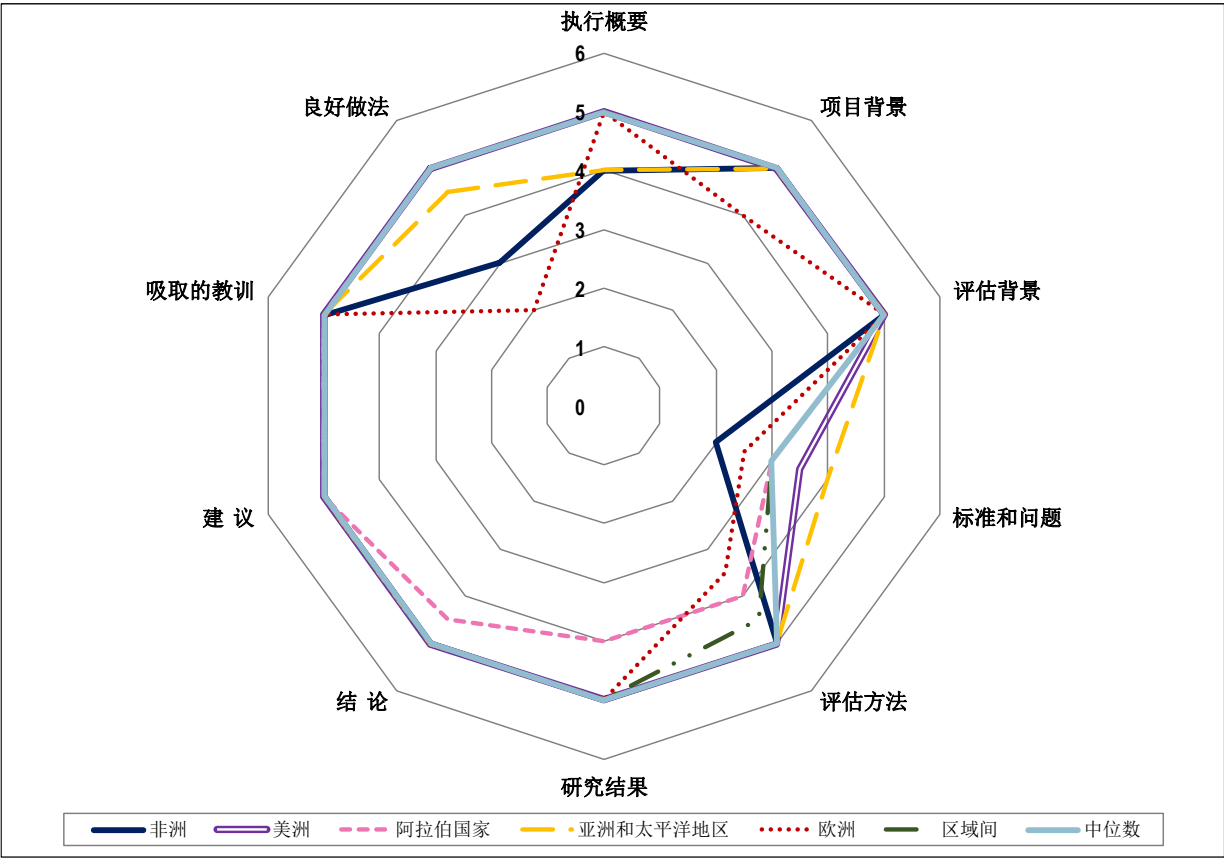
26. 在准则、培训和实践支持方面的投入正在取得成果。从图 3 可以看出，评估办公室通过系统地提高“非常满意”和“满意”的评分，降低“有些不满意”和“不满意”的评分，改善了其结果。值得注意的是，2017-18 年度批次中，没有报告被评为“非常不满意”或“不满意”。
27. 评估办公室已改进进行实时评估的质量评估流程，以确保可以快速确定和解决系统性问题。已据此评估了少量 2019 年的报告。初步结果表明，评估办公室在 2017-18 年度期间实现的质量改进可能难以持续。
28. 质量评估结果还提供了有关质量问题的重要反馈回路，可用于更新评估准则。这种做法涉及修订指导说明，其中包含五个模块，并辅以另外 23 个指导说明，12 个清单和 16 种文书和工具。国际劳工组织第四版评估政策准则预计将于 2019 年底完成，并将于 2020 年初出版。

图 3. 2015-18 年评估报告质量评价



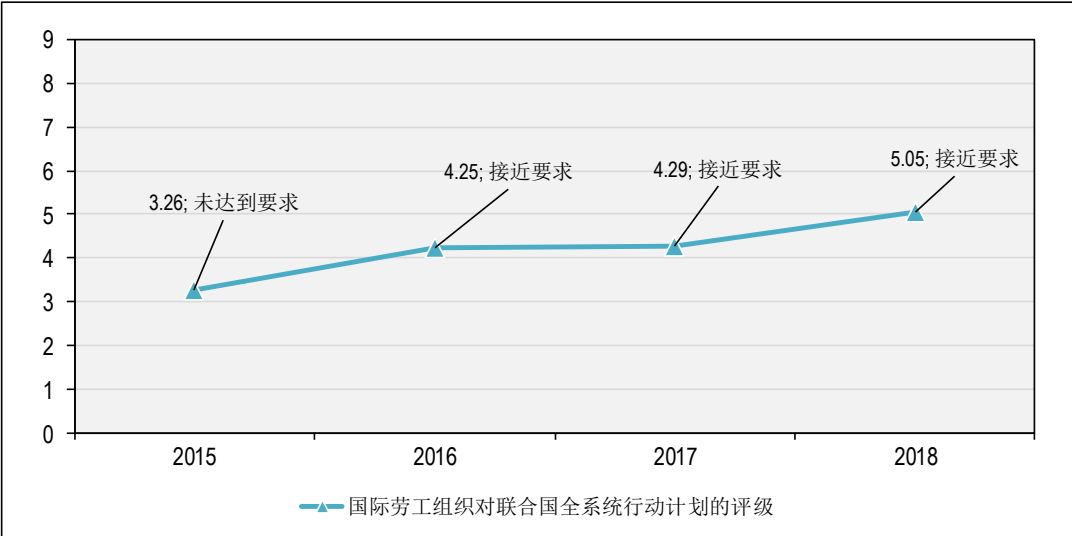
29. 对每个评估项目的中位数结果进行的对比分析表明，各个评价项目之间存在一些有趣的区域间差异(标准、问题和良好做法)。如图 4 所示，总体而言，大多数区域在大多数项目上的评分都令人满意，其中有些区域表现特别出色。需要分析标准和问题方面的表现，以评估如何解决难点问题。

图 4. 各评估项目基于区域划分的评估报告质量评价中位数结果



30. 对《联合国全系统性别平等和增强妇女权能行动计划》(全系统行动计划)平均评级的年度和外部评估比较显示，自 2015 年以来略有稳步改善，表明国际劳工组织正在努力满足全系统行动计划的要求(图 5)。

图 5. 国际劳工组织分散评估显示联合国全系统行动计划评级走向



指标 2.2.2: 通过减少对发展合作项目区域评估的监督，总部评估办公室下放更多权力，以聚力于新的评估模式。	基线: 2016 年评估职能外部独立评估确定将区域一级的独立性问题确定为优先事项，并建议将区域评估员整合为评估办公室的正式工作人员。	状态
两年期里程碑 (2018–19): 编写一份详细的报告分析区域评估员汇报流程，包括各种可能出现的情况，以确保最高水平的独立性。		已达成。
最终目标 (2020–21): 2021 年底前，所有区域评估按照最高独立标准进行，只需要最低程度总部评估办公室监督。		

31. 2016 年对国际劳工组织评估职能的独立评估中确定了评估职能独立性，与此相关的优先事项是，¹¹ 将区域评估员转为评估办公室的正式员工，以加强分散评估的独立性。¹² 评估办公室针对一系列可能出现的情况进行了对比研究。研究结果在联合国其他组织和国际劳工组织内部发现了成功的先例，表明实地工作人员与总部可以是主要上下级关系，与地方主管是次级上下级关系。国际劳工组织工作人员条例对上下级关系的配置做出了选择性规定。如果使用于区域评估官，这会加强分散化评估的独立性。评估办公室计划于 2019 年最后一个季度进一步探讨这个问题。

¹¹ 国际劳工组织：《2011-2016 年对国际劳工组织评估职能的独立评估：最后报告》(日内瓦，2016 年)。

¹² 区域评估官是每个国际劳工组织区域办事处的指定评估官员。他们负责监督各自区域的评估工作。他们对规划程序、管理、体面劳动国别计划审查后续行动以及项目评估进行监督，并提出咨询意见。他们是专门从事支持国际劳工组织评估工作的专业人员。

指标 2.2.3: 组织管理水平评估纳入联合国评估小组规范和标准, 并符合国际劳工组织具体任务及学习需要。	基线: 2013 年独立审查确认质量符合相关要求, 2016 年外部独立评估再次确认。	状态
两年期里程碑 (2018–19): 将更新协议, 在保证质量的同时, 纳入反映国际劳工组织具体任务的新的评估模式。		已达成。
最终目标 (2020–21): 2021 年对国际劳工组织评价职能的独立评价会证实, 组织管理一级的评估根据国际劳工组织具体任务量身定制, 并比照联合国类似机构的类似评估, 仍然具有良好的质量。		

32. 2016 年对国际劳工组织评估职能的独立评估注意到, 国际劳工组织评价程序尚符合国际标准, 但方法需要改进, 以更好地把握国际劳工组织规范性任务。
33. 评估办公室制定了 1 号议定书用以指导针对政策成果和制度评估的组织管理级评估, 还针对《体面劳动国别计划》评估制定了 2 号议定书。根据 2016 年对国际劳工组织评估职能的独立评估的建议, 更新了以上两份议定书, 以反映现有组织战略框架和新的评估政策和战略。此次修订还包括有关性别平等和不歧视的指导意见、国际劳工组织社会对话和规范性任务, 以及国际劳工组织对可持续发展目标的贡献(见子成果 2.4 下的两年期里程碑)。

子成果 2.3 – 为有效政策干预实行知识建设 开展可信度高的影响评估

指标 2.3: 影响评估是可信的, 并用以记录有效政策干预。	基线: 评估办公室 2014 年总结报告指出, 影响评估的质量并非最优也不统一。针对 2018 年将要进行的影响评估样本的事后质量分析将设立新的基线。	状态
两年期里程碑 (2018–19): 提高评估办公室的技术支持水平, 遵守评估办公室对 50% 的影响评估的指导, 从而提高技术部门和国际劳工组织各办事处的影响评估水平。		尚未取得进展。
最终目标 (2020–21): 2021 年底前, 国际劳工组织 85% 的影响评估将具有可信度, 并符合质量及相关性标准的要求。		

34. 评估办公室继续通过“影响评估审查机制”为各部门和区域进行的影响评估提供支持。然而, 因需求不足, 实现两年期里程碑方面进展甚微或者没有进展。原因之一可能是“影响评估审查机制”缺少知名度。另一种可能是, 各单位对将其影响评估的设计和结果交由一个评价性机构审查程序检验能否增加可信度普遍持有疑虑。评估办公室为了帮助他们解决此类困难, 将完成早已在计划中的影响评估事后质量评价。

子成果 2.4 – 更符合国际劳工组织权责和 包括可持续发展目标在内的具体情况的评估框架

指标 2.4: 适用于各级评估的国际劳工组织专用评估方针、模式和方法。	基线: 目前, 极少国际劳工组织专用评估方针及模式用于国际劳工组织评估。	状态 部分达成。
两年期里程碑 (2018–19): 在 5 个试点评估中开发并使用了试点评估框架; 更新了评估政策指导原则。		
最终目标 (2020–21): 更新的评估框架应用于 50%或更多评估, 20%的评估含有可持续发展目标专用指标。		

35. 评估办公室研究了如何改进评估才能使其更好地应对本组织的任务和社会对话机制。第一步, 制定指导说明, 阐述将社会对话和规范性情况系统地纳入项目设计、监测和评估。该指导说明将被纳入试点评估, 以获取更具战略性的方法。2019 年, 为了确保评价人员了解和掌握国际劳工组织的具体情况, 评估办公室制定了一项自我引领计划, 以提高评价人员对国际劳工组织工作及其评价政策和实践的特点的理解。

成果 3: 加强评估结论和建议知识库

子成果 3.1 – 借助 *i-eval Discovery* 应用程序 增强评估信息的可获取性及知名度

指标 3.1: <i>i-eval Discovery</i> 应用程序包含所有规划及已完成的评估, 包括建议、教训及好的做法; 内外部用户经常使用 <i>i-eval Discovery</i> ; <i>i-eval Discovery</i> 被认为是国际劳工组织评估信息的门户。	基线: 根据信息技术司(INFOTEC)提供的数据, 在 2018-19 年度期间, 平均使用量为 2,000。	状态 已达成。
两年期里程碑 (2018–19): 进一步开发及使用 <i>i-eval Discovery</i> 和 <i>i-Track</i> 数据库, 支持有针对性的沟通和使用评估信息(目标: 比基线水平增加 25%)。		
最终目标 (2020–21): 2021 年底前, 作为通往可靠国际劳工组织评估信息的门户, <i>i-eval Discovery</i> 应用程序将在内外部广泛使用。目标: 比基线水平增加 50%。		

36. 为了使评估信息更加方便用户和易于获取, 评估办公室于 2016 年推出了 *i-eval Discovery* 应用程序。这是一个基于网络的交互式应用程序, 公开显示所有计划和完成的评价, 以及相关建议、经验教训、良好做法和总结。自推出以来, *i-eval*

Discovery 作为一种吸收评估结果的创新工具，于内外都获得了广泛的认可，同时也有助于为项目设计、实施和组织学习提供信息。¹³

37. 除了其他筛选选项以外，现在还可以通过捐助者和集群搜索 *i-eval* Discovery 应用程序中的评估。*i-eval* Discovery 应用程序已与国际劳工组织其他平台相连接，¹⁴ 以改进对结果报告的访问。2019 年底前，*i-eval* Discovery 应用程序还将公开提供管理层对独立评估建议的回应，从而向三方成员、员工和捐助者提供全套评估信息。¹⁵ 在此两年期内(2018-19)，访问该应用程序的平均用户数为 2,000。¹⁶

子成果 3.2 – 加强对评估结论的定向交流

指标 3.2: 经修改的通讯战略会更好地将评估结论目标设定为管理层、三方成员以及其他用户。	基线: 2016 年外部独立评估承认已经取得的进展(通讯、分析研究, <i>i-eval</i> Discovery 应用程序), 但要求完善评估结论呈现方式, 以利使用。	状态
两年期里程碑 (2018–19): 制定宣传及品牌重塑战略(目标: 2018), 并与宣传部合作推出此战略(目标: 2019)。		已达成。
最终目标 (2020–21): 2021 年针对国际劳工组织评估职能的独立评估会对宣传战略方面取得的进展予以承认。		

38. 评估办公室制定了 2018-21 年宣传计划，以努力深化国际劳工组织的评估文化。根据《2018-21 年国际劳工组织评估战略》，宣传计划旨在让评估结论更好地面向管理层、三方成员和其他用户。
39. 新计划的出台恰逢其时，因为自 2016 年以来，评估办公室扩大了宣传产品和服务的数量，这些产品和服务需要针对及满足利益攸关方需要的宣传计划。已与宣传和新闻部协商，确定能在多大程度上支持推出计划。
40. 过去一年中，评估办公室应用了新的工具以更好地宣传评估信息(图 6)。

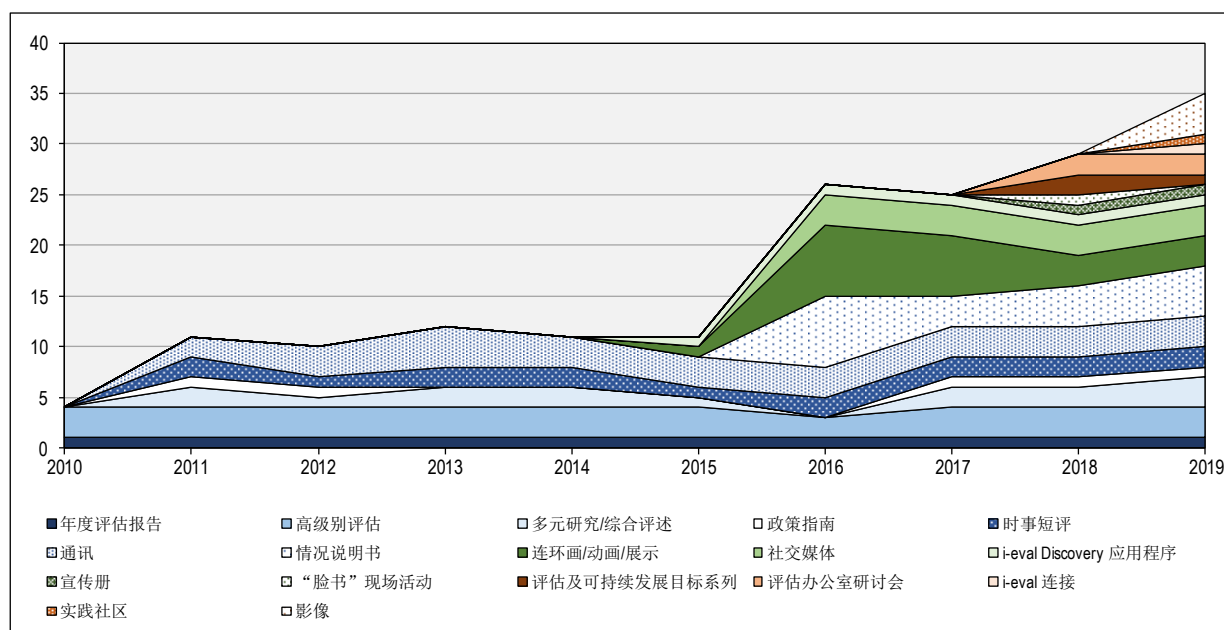
¹³ 这包括超过 1,000 项已完成的评估，5,000 多项建议、1,700 条教训，700 种好的做法以及近 400 项已列入计划的评估(截至 2019 年 6 月)。

¹⁴ [发展合作看板](#)和[体面劳动结果看板](#)。

¹⁵ 2020 年，将更新 *i-eval* Discovery 应用程序底层数据库，以更好地获得评估信息。

¹⁶ 截至 2019 年 5 月，根据 INFOTECH 提供的数据，*i-eval* Discovery 应用程序的年平均使用量在 2,000 用户上下。评估办公室对此有信心 2020-21 两年期结束前增加 50%的最终目标触手可及。

图 6. 2010-19 年评估办公室的产品和服务



子成果 3.3 – 三方成员及管理层更好地 将评估结论及建议用于管理及决策过程

指标 3.3.1: 评估咨询委员会对评估的时间安排和利用提出的建议促使更好地吸收评估结论，以利制定全球及区域层面的政策及决策。	基线: 评估咨询委员会平均每年召开四次会议，定性分析表明，委员会就 40% 组织管理层评估进行了战略辩论。尽管各区域都加入了评估咨询委员会，但尚未有形成区域评估咨询委员会。	状态
两年期里程碑 (2018–19): 评估咨询委员会继续召开会议(每年四次)，就 50% 组织管理的评估进行战略讨论。2019 年初前，将撰写一份关于区域咨询委员会附加值的报告。根据该项审查的结果，2019 年底前，将有两个区域试行区域评估咨询委员会。		部分达成。
最终目标 (2020–21): 2021 年底前，评估咨询委员会继续召开会议(每年四次)，就 75% 组织管理层评估进行战略讨论，仍继续重视联合支持，以解决评估中发现的系统性问题。目标是推广将建立的区域评估咨询委员会的实践做法，具体视试点结果而定。		

41. 2018-19 两年期里程碑要求每年召集四次会议，并就 50% 组织管理层评估进行战略讨论。¹⁷ 在 2017-18 年度评估报告中，在报告期所审查的组织管理层评估中，有 75% 属于战略讨论的主题事项。除此以外，所有已列入计划的会议中，已召开了三次会议。

¹⁷ 本里程碑取决于评估咨询委员会要求修改的工作计划或对被认为不符合要求或不完整的进度报告要求递交二次口头报告的数量。

42. 表 5 显示评估咨询委员会当前报告所述期间决定的范围。2019 年最早的两次会议上，委员会针对组织管理层评估召开了五次讨论会，完成并批准了工作计划。因此，五次讨论会都应是战略性的。
43. 2018 年和 2019 年的数字相加(即九次讨论会，其中八次为战略性的)达到 89%，远远超过了 50%的目标。虽然已达到两年期里程碑，但评估办公室认识到其计算方法有一定局限性。下一个两年期内，评估办公室将与评估咨询委员会一起探索新的计算方法。年底前评估办公室将开展研究探索建立区域评估咨询委员会的附加值。

表 5. 2019 年评估咨询委员会关于高级别评估的决定

高级别评估	工作计划状态	回顾评估咨询委员会召开的讨论会，包括关于工作计划和进展报告的意见
湄公河区域 (2017)	已完成 ¹	2018 年 2 月，工作计划获得批准，并要求提交进度报告。日程安排出现一些延误，口头进度报告于 2019 年 2 月获得批准。
外地业务和结构(2017)	已完成	2018 年 2 月，后续工作计划被退回修订。2018 年 5 月，主管外地业务和伙伴关系的副局长再次被邀请修订和重新提交。2018 年 10 月，评估咨询委员会批准了经修订的后续工作计划，并要求在六个月内提交进展报告。2019 年 5 月，进展报告获得批准。讨论的重点是，评估结论需要为劳工局时刻准备快速应对联合国改革外地结构可能产生的影响提供资讯。
青年就业 (2018)	已批准 ²	2019 年 2 月，评估咨询委员会批准了评估建议后续工作计划，要求于 2019 年第三季度提交口头进展报告。
阿拉伯国家 (2018)	已批准	2019 年 2 月，评估咨询委员会批准了后续工作计划，确认六个月后提交一份关于其执行情况的口头进展报告。
能力建设 (2018)	已批准	2019 年 2 月，评估咨询委员会批准了稍作修改的工作计划，并确认六个月后提交关于其执行情况的口头进展报告。战略讨论涉及需要进一步细化能力发展机构战略。

¹ “完成”表示评估咨询委员会已经接受了有关责任部门提交的口头和书面进展报告定稿，没有要求再提交其他报告。

² “批准”只代表批准工作计划。

指标 3.3.2: 通过系统监测加强对评估建议的后续行动。	基线: 2016 年，管理层回应后续行动为 83%(部分处理并完成)。	状态 部分达成。
	两年期里程碑 (2018–19): 将建立管理层对评估建议后续行动的自动在线申请，提高整体效率并保持后续行动高回应率(目标: 85%)。	
	最终目标 (2020–21): 2021 年底前，管理层评估建议后续行动的自动在线申请将提高评估质量，以及管理层回应评估建议的质量(目标 90%)。	

44. 管理层对评估建议的回应对于确保评估的使用和维护报告的可信度至关重要。国际劳工组织要求各级管理人员对建议的后续行动负责，指出建议是否已经完成或部分

完成，是否未计划采取行动，¹⁸ 是否尚未采取行动，¹⁹ 建议是否被拒绝，或者建议是否不适用(表 6)。

表 6. 2018 年管理层对评估建议的回应

行政区域 或部门	要求管理层回应的评 估报告(52)		已完成	已部分 完成	尚未计 划采取 行动	尚未采取 行动	拒绝	不适用
	建议数量							
	报告中	附有管理层 回应						
非洲	146	146	66	38	15	24	3	0
阿拉伯国家	91	91	38	30	17	5	1	0
亚洲	85	85	3	54	1	15	0	12
欧洲	6	6	0	6	0	0	0	0
拉丁美洲和加勒比	58	58	13	30	5	10	0	0
区域小计	386	386	120	158	38	54	4	12
更好的工作处	8	8	8	0	0	0	0	0
就业政策司	24	24	16	6	1	1	0	0
工作中的基本原则 和权利处	26	26	1	16	6	2	1	0
GED ¹	7	7	6	0	0	0	1	0
劳动行政管理和 监察计划	9	9	0	8	0	0	1	0
SFP ²	2	2	1	1	0	0	0	0
办事处/单位小计	76	76	32	31	7	3	3	0
累计	462	462	152	189	45	57	7	12
百分比	0	100	32.9	40.9	9.7	12.3	1.5	2.6
已完成、已部分完 成管理层回应总数			73.8%					

¹ 性别、平等和多样性及艾滋病毒、艾滋病和劳动世界处。

² 社会融资计划。

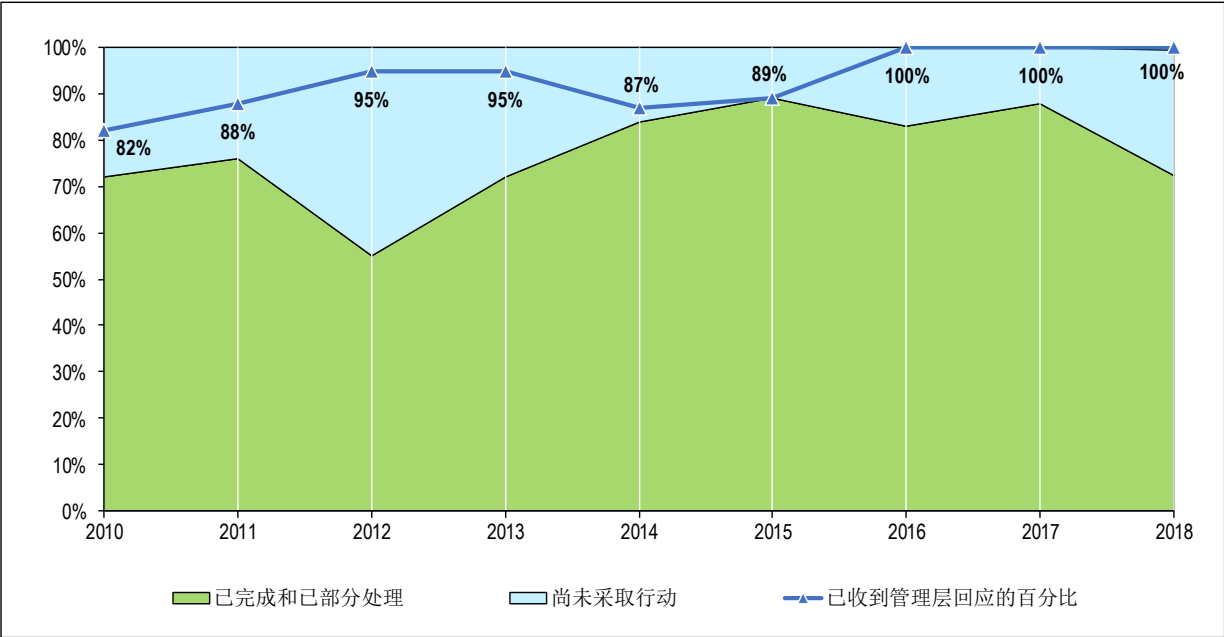
45. 审查所涉期间(2018 年)，从要求跟进的独立项目评估那里收到了所要求的全部 52 份管理层回复，代表对 462 项建议的管理层回复。总共已完成或部分完成 74%的建议，低于 2017 年的 88%(图 7)。平均占 81%，因此没有达到 85%的两年期目标。

¹⁸ “未计划采取行动”指已采纳建议，但未计划针对建议采取行动。

¹⁹ “尚未采取行动”指已采纳建议，但尚未实施针对建议的行动。

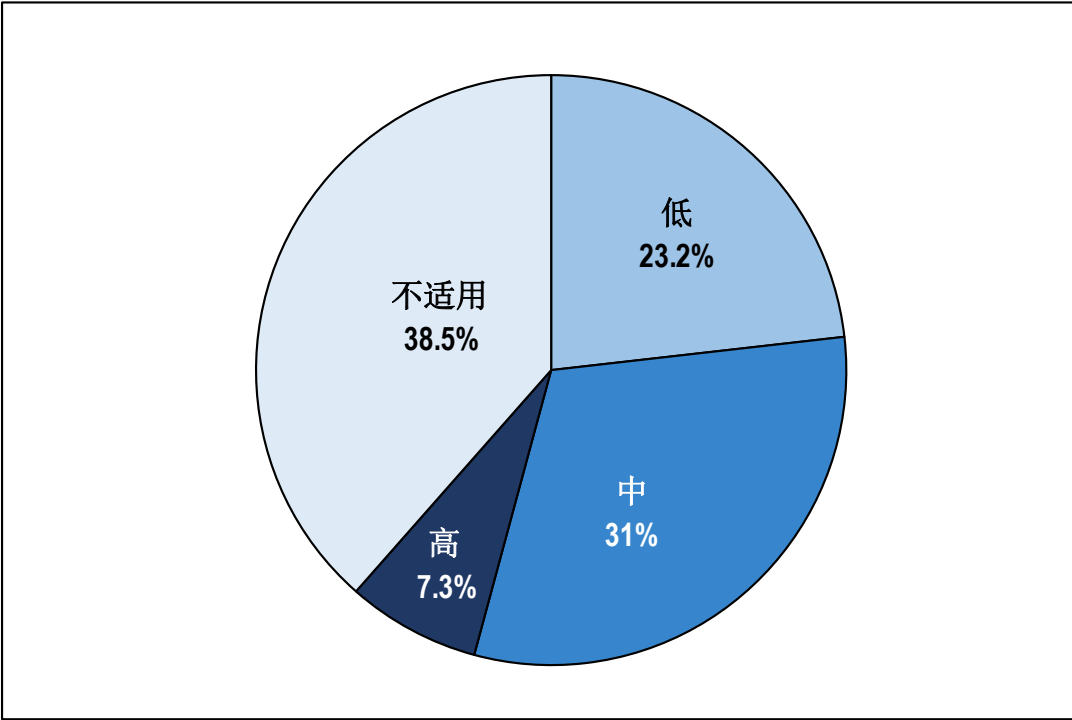
46. 管理层对评估建议的回应跟进下降有两个原因。首先，修改了管理层回应的后续行动方法，以便更严格地评估对建议的后续行动的质量。其次，跟踪管理层回应的新的自动化在线系统²⁰与以前由评估办公室手动管理的系统有所不同。新系统使管理人员能够直接跟踪评价建议，同时也允许自动整合所有数据，以满足实时报告要求，从而提高效率。
47. 分析表明，大多数建议并不涉及大量资源影响，因此表明费用并不是采取行动的障碍(图 8)。此外，将近 77%的建议针对国际劳工组织各部门和外地办事处；几乎 23% 与国际劳工组织的三方成员相关。

图 7. 2010-18 年对评估建议的后续行动的进程



²⁰ 2018 年 2 月启动了自动管理响应系统。

图 8. 2018 年按资源影响划分的建议百分比

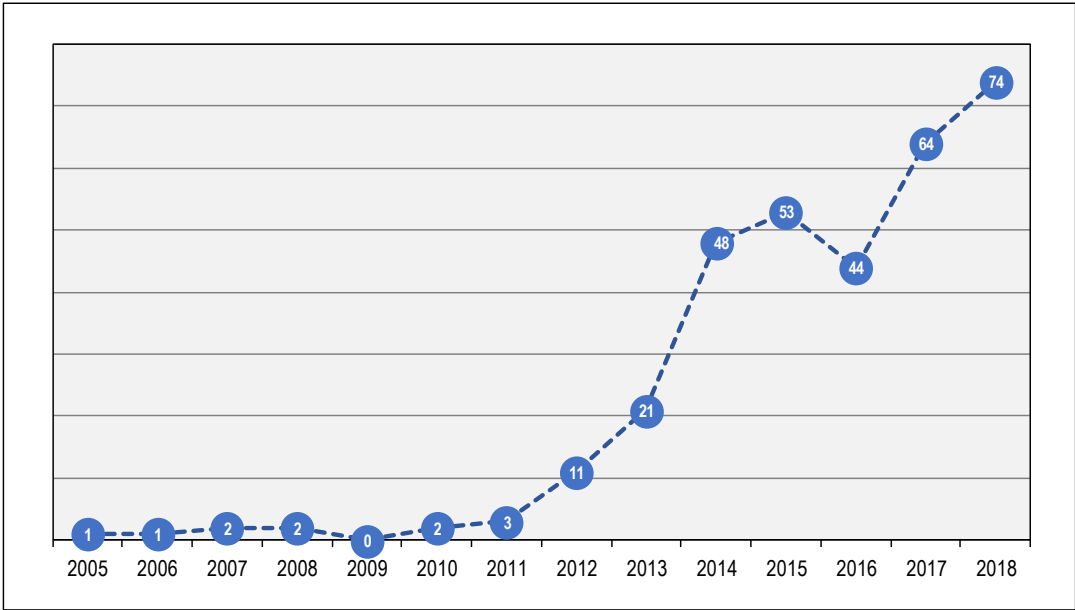


指标 3.3.3: 加强在报告战略指导、审查和战略规划、计划和预算记录以及其他高级别规划和战略中使用评估。	基线: 年度评估报告记录采用建议和从评估中吸取的教训(2010-15 年 40-50%，基于总结)。	状态
两年期里程碑 (2018-19): 2019 年底前，75%的评估建议和结论全部或部分反映在相关的战略指导和报告中(例如，实施报告、2020-21 年计划和预算报告和其他战略和规划文件)。		部分达成。
最终目标 (2020-21): 2021 年底前，80%的评估建议和结论全部或部分反映在相关的战略指导和报告中(例如实施报告，2020-21 年计划和预算报告以及其他战略和规划文件)。		

48. 以上子成果旨在对评估建议和结论在决策和实践中得到反映的程度进行评价。这是一个雄心勃勃的目标，很难评价。但是，可以研究许多替代措施。例如，理事会所讨论的高级别评估提出的建议中有 82%已被完全采纳，并在本报告所述期间采取了行动。
49. 获取评估结论有时可能是漫长而困难的过程。越来越多的人认识到需要纳入明确的变革理论，这说明评估结果对实践具有影响(图 9)。2010 年以来，国际劳工组织的评估报告越来越多提到这一问题，以至于其重要性作为一项建议被纳入 2013-14 年度评

估报告，²¹并在随后的年度评估报告中继续得到强调。为了进一步提高变革理论的重要性，内部基于结果的管理工作队在 2018 年将其确定为优先事项之一。

图 9. 国际劳工组织公共网站上包含“变革理论”一词的文件



来源：国际劳工组织国际培训中心培训资料“创建基于结果的变革理论”。

50. 通过对监测和评估框架优缺点进行评价，国际劳工组织工作的可评估性有所提高。理事会第331届会议(2017年10月至11月)批准了《2016-17年度评估报告》中的建议后，所有高价值发展合作项目现在都需要进行可评估性评价。此外，2017年，评估办公室开发了一项诊断工具，以提高《体面劳动国别计划》的可评估性，并有利于报告国家计划目标及其对可持续发展目标的贡献。在此经验的基础上，计划司同意在《体面劳动国别计划》的编制中纳入可评估性内容的要求(见子成果 1.4 下的两年期里程碑)。
51. 对国际劳工组织能力建设(2018 年)的高级别评估的所提建议，有 88%明确提及在一份或以上全球和国家级战略文件中的现在或将来针对该建议的解决方式。理事会第 335 届会议(2019 年 3 月)批准的国际劳工组织全系统能力发展机构战略²²多次提到并使用了这项评估。
52. 作为系统程序进程的一部分进行的一项非详尽调查表明越来越多高级别评估所提建议和结论被采用，并在各个区域发挥作用，由此，包括在联合国框架内，致使在评估中确定新的研究课题并使战略立场的形成有据可依。此程序系统用于记录涵盖评估咨询委员会的讨论以及自动化管理层回应系统已有体系之外的后续行动。

²¹ 多边组织业绩评估网(MOPAN)以及外聘审计师分别在 2016 和 2017 年呼吁在组织层面更好地阐述变革理论。

²² 理事会文件 [GB.335/INS/9](#)。

子成果 3.4 – 通过进一步分析评估结论及结果， 利用评估满足战略知识的要求

指标 3.4.1: 分析、综合评估结论并将其录入知识成果，支持计划制定及知识建设。	基线: 上一个战略阶段期间，撰写了 22 篇研究分析文章，进行了多元研究和综合审查。	状态 部分达成。
两年期里程碑 (2018–19): 设立了符合战略知识要求的专题确定过程，保持每年平均至少进行三项研究。		
最终目标 (2020–21): 2021 年底前，所产生的知识项目数量将已增加 25%，2021 年对国际劳工组织评估职能的独立评估将确认符合战略知识要求的专题。		

53. 评估办公室对分析和汇总符合战略知识要求的评估结果的需求作出回应。自 2011 年以来，根据技术部门的要求进行了综合审查、多元研究，撰写了分析文章，而其他部门则为国际劳工大会的周期性讨论提供支持，或回应评估办公室对国际劳工组织整体有效性进行报告的要求。本两年期期间，评估办公室完成了 9 项综合审查、多元研究和分析文章，与每年三项研究的基准相比，每年平均进行 4.5 项研究。
54. 虽然评估办公室在确定战略知识需求议题时反应迅速并具有远见，但建立这类议题的正式程序尚未全部完成。评估办公室将使 2018 年承担的理事会文件和其他报告的内容分析正规化，并通过持续对话和有条理的调查对其进行补充，以更好地了解三方成员评估方面的需求。初步调查表明，劳动改革、青年就业、移徙、和平与复原力以及性别主流化等仍然是典型的议题。2019 年底前，评估办公室将在评估网络之外进行协商，以建立一个更有效的机制来确定评估需求。

指标 3.4.2: 年度评估报告提供国际劳工组织总体有效性的年度概况。	基线: 完成对体面劳动结果及国际劳工组织发展合作有效性的分析，覆盖 2009-16 年度，同时修订方法。	状态 已达成。
两年期里程碑 (2018–19): 对 2017 年和 2018 年进行分析，提供关于国际劳工组织有效性的综合报告；进一步改进方法，以便在年度评估报告中定期进行分析和报告。		
最终目标 (2020–21): 分析进行到 2021 年，并传达给国际劳工组织的相关部门以供使用，年度评估报告负责报告对评估结论的吸收和使用。		

55. 2011 年以来，评估办公室每年都对国际劳工组织的战略相关性、有效性、影响、可持续性和效率进行分析。《2018-21 年评估战略》要求通过指标 3.4.2 继续进行分析。因此，本报告第二部分在已改进的方法的基础上，介绍国际劳工组织绩效成果。

第二部分. 国际劳工组织有效性及结果评估

56. 本报告第二部分概述评估办公室在基于结果的管理方面的举措。通过验证结果并鼓励整合汲取的经验教训与管理决策，评估有效性和问责制。

1. 制定国际劳工组织绩效在国家及全球两级更全面的验证程序：进展受限的内外部因素

57. 国际劳工组织最近的政策和管理文件²³充分反应人们对有效性评估和测量的兴趣日益浓厚，这是获取体面劳动结果和效力的优质、可获得、及时且可靠信息的手段。新的评估政策和战略以及劳工局内部基于结果的管理工作队正在进行的工作催生了雄心勃勃的改革计划。评估办公室致力于制定变革性方针，以制定更全面的国际劳工组织国家及全球两级绩效验证程序。理事会在去年的年度评估报告中提出并认可了新的评估设想，它们旨在过渡到一个通过集群评估全面覆盖所有国家计划(《体面劳动国别计划》)以及计划和预算成果的系统。

58. 在本报告所述期间，尽管慢于预期，在实现以上目标方面已有进展。已就如何使评估框架、方法和程序符合国际劳工组织标准和社会对话任务以及背景编写了新的指南。尽管在推进集群评估和审查区域评估干事的上下级关系方面有所延误，但关于集群评估的原则、程序和做法的内部研究已经完成，根据一整套归类标准确定了试点项目(见第一部分，成果 2)，并且 *i-eval Discovery* 应用程序已经能够发挥集群评估规划工具的作用。需要加强对战略性集群评估和募集评估准备金的内部和外部支持(见建议 1)，以遏制继续倾向于支持条块分割、针对具体项目的方法。

2. 国际劳工组织工作的整体有效性

59. 2011 年以来，评估办公室已进行了一系列的发展合作项目评估两年期多元研究。²⁴ 最新研究通过审查 2017-18 年独立项目评估，对国际劳工组织体面劳动结果进行公正评价。评估办公室还对 2013 至 2017 年间经常预算补充账户出资的干预措施的有效性进行了重点可比研究。2011 年以来，用于这些研究的方法和途径仍具有相对可比性。用于评估的 26 个绩效指标中的每一个指标都按照 4 分制进行评级：不成功(1)；部分成功(2)；成功(3)；非常成功(4)。

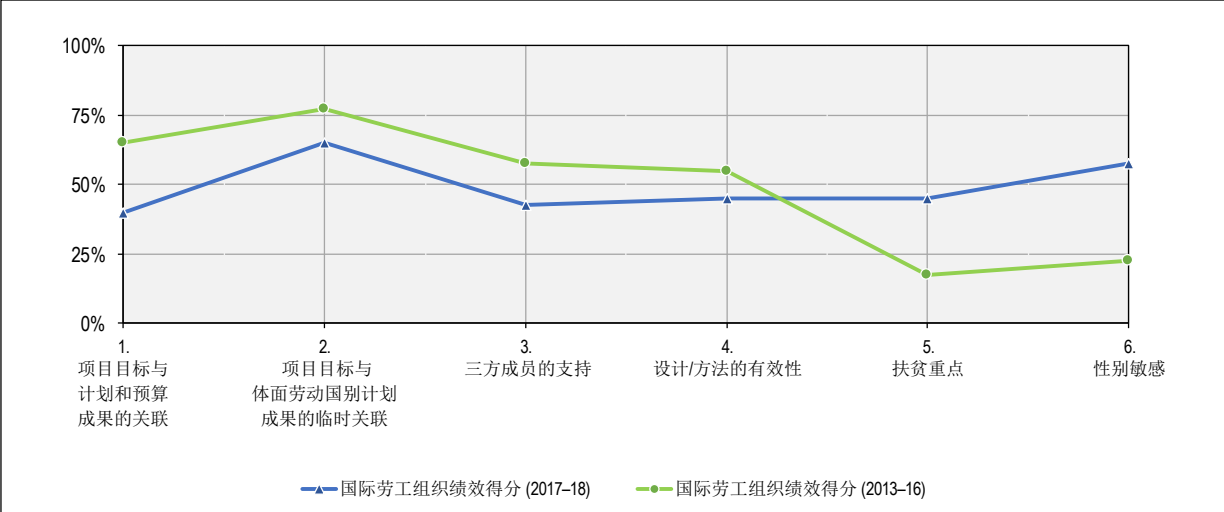
²³ 例如，2018 年国际劳工大会关于开展有效发展合作以支持可持续发展目标的决议再次要求“更好地利用数据采集和包括影响评估在内的基于结果的管理工具及评价，以显示其中有效的部分；扩大干预措施，提高《体面劳动议程》的知名度”。

²⁴ 国际劳工组织：体面劳动结果以及国际劳工组织技术合作有效性：2009-2010 年项目评估多元分析(日内瓦，2011)；国际劳工组织：体面劳动结果以及国际劳工组织技术合作有效性：2011-2012 年项目评估多元分析(日内瓦，2013)；国际劳工组织：体面劳动结果以及国际劳工组织技术合作有效性：2013-2016 年发展合作多元分析(日内瓦，2017)。

60. 整体而言，国际劳工组织2017-18年间已评估的发展合作项目的有效性、可持续性和影响方面的评级为“成功”，而战略相关性、一致性、管理和资源的实施和效益方面的评级仅为“部分成功”。

2.1. 战略相关性和一致性

图 10. 战略相关性和一致性方面获得“非常成功”或“成功”评级项目的百分比

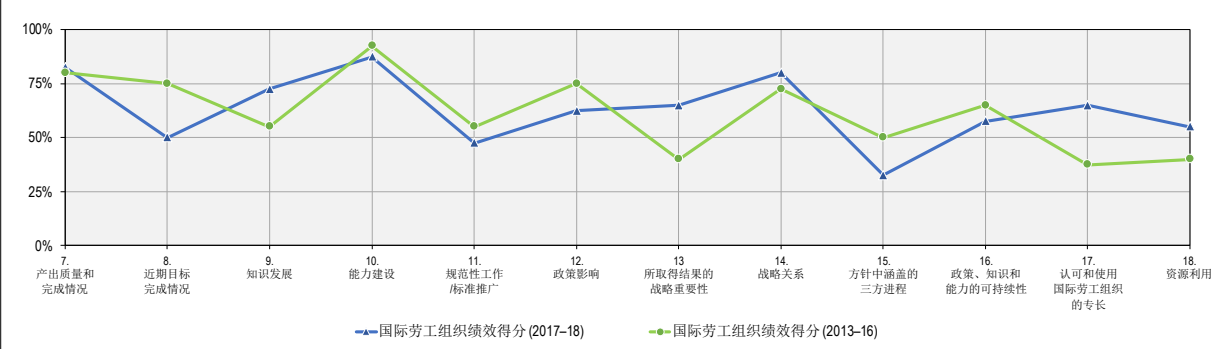


注释：每个绩效标准的百分比对应于获得 3 或 4 等级项目的比率。

61. 最高得分(见图 10)出现在评估发展合作项目目标与《体面劳动国别计划》或国家计划成果的相关性(标准 2)，以及发展合作项目在多大程度上包括性别敏感指标以及解决性别不平等问题总体战略(标准 6)。与项目设计质量相关的绩效领域得分较低，特别是在三方成员的参与和设计的有效性方面(标准 1、3 和 4)。令人惊讶的是，项目设计和实施中对减贫工作的处理方面得分仍然很低，只有 5%的项目明确说明贫困影响或采取扶贫措施(标准 5)，获得“非常成功”的评分。

2.2. 有效性、可持续性和影响

图 11. 有效性、可持续性和影响方面得到“非常成功”或“成功”评分项目的百分比



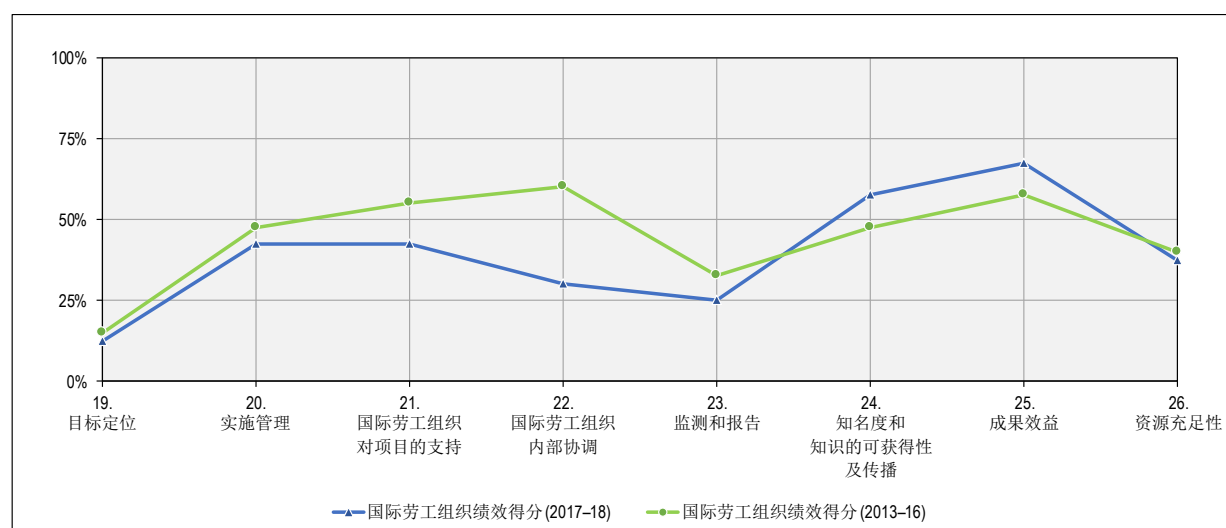
注释：每个绩效标准的百分比对应于获得 3 或 4 等级项目的比率。

62. 2017-18 年期间，发展合作项目的有效性、可持续性和影响方面的总体表现(见上文图 11)仍然良好。主要优势领域包括产出的质量和完成度(标准 7)以及个人和机构层面的能力建设(标准 10)，其次是利用战略关系和发展知识的有效性(标准 14 和 9)。国际劳工组织的专家建议经常得到承认和采用(标准 17)，半数以上的项目因其政策影响和结果所具有的战略重要性而引起注意(标准 12 和 13)。

63. 据观察，项目利用资源和为长期保持成果创造条件的能力表现参差不齐(分别为标准 18 和 16)。此重要类别中的主要弱点在于对相关劳工标准的促进和纳入三方性(标准 11 和 15)，鉴于国际劳工组织在规范和社会对话方面的特定权责，这一点至关重要。

2.3. 管理和资源的实施及效益

图 12. 管理和资源的实施及效益方面得到“非常成功”或“成功”评分项目的百分比



注释：每个绩效标准的百分比对应于获得 3 或 4 等级项目的比率。

64. 实施管理及其效益的绩效令人关切，大多数项目在总体绩效标准上得到的评分大多较低或一般。国际劳工组织项目的成本效益(标准 25)除外，因为它在几乎三分之二项目中的评分都很高。知识和传播的知名度和可获得性(标准 24)得分很高；然而，国际劳工组织对项目实施的支持、内部协调以及人力和财政资源的充分度仅被评为部分成功(标准 21、22、26)，反应出有很大改进余地，这应该不难做到。不到半数项目的项目管理得到肯定性的评分。

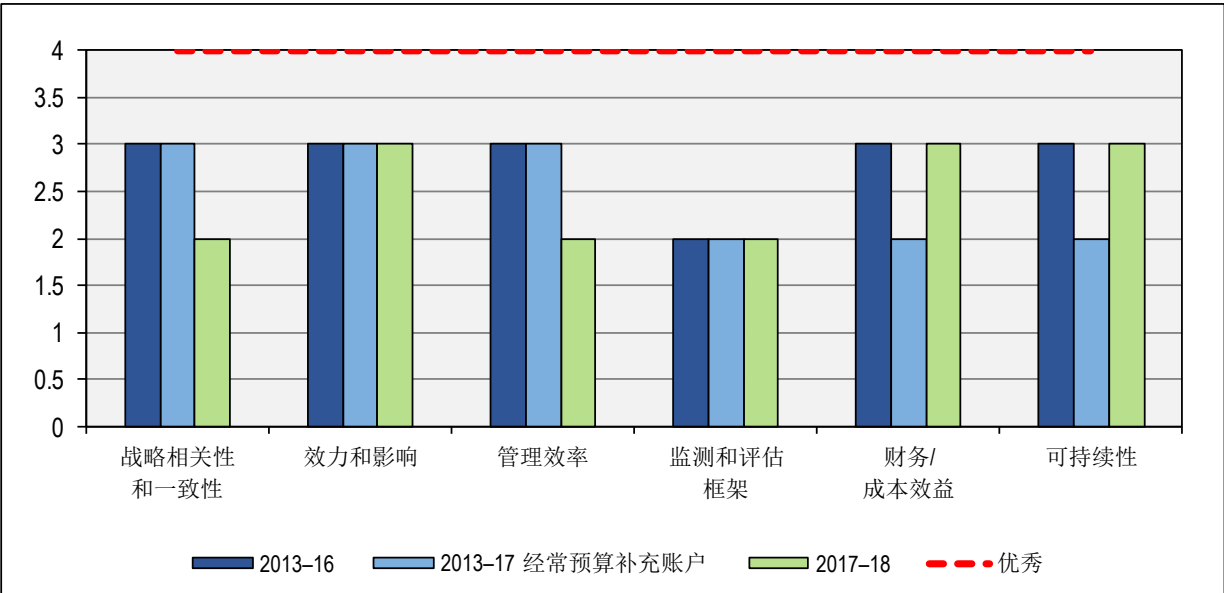
65. 在监测和报告方面，只有四分之一的项目被评为“成功”(标准 23)。监测框架(存在时)存在资源不足的问题，且可评价性弱。报告的重点往往是活动一级取得的进展，而不是总体目标。

66. 此类别中得分最低的是目标定位(标准 19)，这表明设计基于结果的监测框架一直是难点。

2.4. 国际劳工组织 2013 年以来的绩效

67. 一段时间的绩效结果比较显示，无论年份和资金来源，国际劳工组织项目有效性具有整体实力，而在项目的目标定位、监测和评估框架、采取扶贫政策、管理效率和资源充足性方面经常存在弱点(图 13)。

图 13. 2013-18 年国际劳工组织整体绩效得分



注释：该图描述国际劳工组织绩效的趋势，用每项绩效标准的总得分的中位数表示。在此次多元研究中，首次使用中位数代替平均值来衡量中心趋势，以更好地说明分配给绩效标准的最常出现的分数，因为中位数比平均值受极高或极低值的影响更小。最近还将中位数引入项目评估的事后质量评价方法(见子成果 2.2.1)，并将继续将其应用于即将进行的体面劳动结果滚动多元分析。

68. 在项目设计合理程度、与三方成员关系密切程度、促进规范性工作以及国际劳工组织在实施过程中的内部协调方面，2017-18 年开展的发展合作项目的绩效低于 2013-16 年。已观察到以下方面的绩效有所提高：项目产生知识的能力、采用国际劳工组织专家建议以及利用资源提高项目成果的能力。2017-18 两年期期间，对性别问题的回应得到了很大提高，近 60% 的项目显示出较高覆盖率。2013 年以来，共同的优势领域是获取和利用项目产出，能力发展以及伙伴关系的战略建设。

2.5. 已评估的预算外发展合作供资项目样本的有效性

69. 将经常预算补充账户供资的发展合作项目与常规发展合作项目进行对比，注意到以下方面的绩效比较突出：设计的合理性(80%优良绩效得分是因其符合计划和预算以及《体面劳动国别计划》成果)；国际劳工组织对项目实施给予支持；已取得的结果具有战略重要性。其绩效在以下方面低于平均：监测和评估框架的质量(所有项目的

绩效得分都较低)；²⁵ 近期目标的完成率和可持续性；对性别问题的回应；以扶贫为重点的干预措施。共同优势反映在：在知识发展、能力建设和政策影响等重要领域的产出的技术质量和取得的结果。

3. 促进组织学习：有利于成功项目效力的主要驱动因素

70. 对评估的多元分析所确认的成功的重要驱动因素在于项目是否将其战略目标与《体面劳动国别计划》以及国家计划成果紧密联系(显示高关联度)，在于项目在多大程度上涉及国际劳工组织三方成员，并在起始阶段就推行三方性(由此获得高相关度的产出和对项目的把控能力)。
71. 有利于有效取得成果的其他经常性驱动因素包括利用内部和外部伙伴关系以及建立和实施运作良好的项目管理程序。成功的干预措施全面完成了其计划产出，所有这些产出都有较高的技术质量，还通过强有力的知识传播策略显示出较高知名度。先前的多元研究已强调上述因素，而当前的发展合作项目评估分析则确认了其他成功的驱动因素。²⁶ 人力和财政资源达到所需水平以及与其他项目和技术专家进行战略合作，都是确保物有所值的关键所在。
72. 完善的项目设计与所获结果被三方成员使用之间存在正向且统计上显著的关联。相关结果能否长期切实得到维持，总的来说取决于设计得当的监测和评估框架，还取决于结果的相关性的增加和三方成员对结果的使用。与此相应，利益攸关方认为对其国家发展结果具有高度战略意义的项目，与国际劳工组织在实现项目近期目标(包括解决性别不平等问题战略)方面取得更大成功之间，存在紧密联系。
73. 人们认为结果具有战略重要性，往往与它们对促进和适用劳工标准的潜在贡献有关，还与三方成员将结果用作依据以在国家一级获得影响政策的机会有关。

建议 2：为提高国际劳工组织工作的整体有效性，考虑到本报告确定的经常性成功驱动因素，并一贯将这些因素纳入项目设计和执行。

²⁵ 这主要因为没有针对经常预算补充账户支持的项目和国家计划成果的正规监测和报告机制，由此阻碍劳工局记录结果进展的能力。

²⁶ 影响因素基于多元研究的发现，使用统计相关分析对影响因素进行了三角测量，这些统计确定了中等到强的正相关和统计上显著的相关系数。应该注意的是，两个变量之间的相关性表明了某种程度的关系，并不一定存在因果关系。

决定草案

74. 理事会核可《2018-19 年度评估报告》提出的由国际劳工组织执行的建议 (理事会文件 GB.337/PFA/6 号第 22 和 73 段)和 2019-21 年评估工作计划的优先事项。

附录. 执行《2017-18 年度评估报告》中所载已批准建议的行动计划

建议	长期改进 (到 2021 年或按照 确定的时间表)	短期行动 2018-19 年	部门/额外费用	状态	评估办公室意见
1. 加强国际劳工组织基于结果的框架					
建议 1: 将好的做法确定下来, 进度报告定稿包括自我评估部分, 而不再需要单独的正式自我评估报告。	增加与评估要求的符合度; 减少项目负责人的报告负担; 改进组织学习。	审查进度报告定稿的格式 进度报告定稿中包括自我评估部分 监测进展	伙伴关系与外地支持司(PARDEV)/ 评估办公室 (无费用)	伙伴关系与外地支持司与评估办公室针对发展合作进度报告定稿开展合作, 该报告纳入自我评估部分(D 部分)。它适用于高达 50 万美元的所有发展合作项目。	伙伴关系与外地支持司现在没有跟踪符合自我评估要求的进度报告定稿的数量。因此劳工局就无法评价与自我评估要求的符合度。
建议 2: 改进全劳工局的监测和报告框架, 以及预算外资助活动的做法, 明确问责制。	劳工局内部明确的问责制框架, 以进行劳工局范围内的综合项目监测和报告, 从而提高绩效数据的可用性和一致性, 并提高评估报告的质量。	审查问责制框架, 以进行进度监控和报告 确保整合劳工局整体结果监测和报告	伙伴关系和外地支持司与计划司合作(未提供费用)	第 154 号内部管理文件规定了预算外发展合作(XBDC)供资项目的监测和报告问责框架。国际劳工组织负责官员负责项目监测以及及时编写和提交高质量的捐助者报告。自 2018 年引入捐助者报告在线平台以来, 预算外发展合作供资项目的报告做法有所改进, 该平台为捐助者报告的中央储存库和管理工具。这提高了数据的可用性。报告通常由国际劳工组织负责官员直接提交给捐助方。伙伴关系和外地支持司无权也没有责任对技术进度报告进行质量保证。 作为《2018-19 年计划执行报告》的一部分, “体面劳动结果看板”的修订版包括一个新的功能, 以显示预算外发展合作供资项目对整体结果的贡献, 促进与“发展合作看板”的联系。就全劳工	2010 年发布的第 154 号内部管理文件先于理事会 2018 年认可的新评价政策和相关建议。提出这项建议的依据是在提交、质量控制和储存进展报告方面发现了弱点, 从而在发展合作报告的问责制框架中留下了很大缺口。这些问题应在劳工局将建立的组织监控系统中解决, 以确保问责制并促进 2020-21 年的报告。

建议	长期改进 (到 2021 年或按照 确定的时间表)	短期行动 2018–19 年	部门/额外费用	状态	评估办公室意见
----	---------------------------------	-------------------	---------	----	---------

局范围的监测和报告而言，计划司正在牵头开展三个重要的持续进程：

1. 国际劳工组织基于结果的管理工作队进行了对劳工局内部监测工具和系统的梳理，查明可以精简的空间和潜力。
2. 基于结果的管理工作队还提出了关于应用变革理论方法的具体建议，从而修订了撰写《2020-21 年工作计划》和结果框架的方法，并制定了一套产出、成果和影响指标，以促进对绩效和体面劳动趋势的监测。
3. 劳工局正在制定一个程序，以进一步加强其系统，提高提供有关资源和结果的信息的透明度。基于这三个进程，劳工局将加强其组织监测系统，以确保问责制并促进 2020-21 年的报告。